心理学报 2020, Vol. 52, No. 7, 847−860
Acta Psychologica Sinica DOI: 10.3724/SP.J.1041.2020.00847

音乐训练促进诗句韵律整合加工的神经过程*

张政华[1,2] 韩 梅[1,2] 张 放[3] 李卫君[1,2]
([1]辽宁师范大学脑与认知神经科学研究中心; [2]辽宁省脑与认知神经科学重点实验室; [3]辽宁师范大学音乐学院, 大连 116029)

摘 要 本研究采用ERP技术, 考察音乐训练组和对照组完成诗句押韵判断任务时, 在绝句末对韵律信息(含声调和韵母两个维度)的整合加工过程。结果发现, 在100~300 ms, 仅音乐训练组在声调/韵母合适条件下, 对韵母/声调的一致性进行深入分析, 并诱发了更大的正波; 在韵母违反条件下, 声调违反相比声调合适诱发了更小的正波。在300~750 ms, 两组被试均在绝句末对诗句内出现的韵母和声调违反进行整合分析并诱发了广泛分布的负波。不过, 对照组仅在声调/韵母合适条件下进行, 而音乐训练组则在声调/韵母违反条件下完成此过程。综上, 音乐训练组和对照组均会在诗句末完成押韵信息的整合加工, 但是音乐训练组对韵律信息(尤其是声调)的加工更敏感和快速, 并且对不同类型的违反有更精细的差异性反应

关键词 音乐训练; 韵律; 诗句; 整合加工
分类号 B842

1 引言

长期的音乐训练会引起皮层(Chen, Penhune, & Zatorre, 2008; James et al., 2014; Steele, Bailey, Zatorre, & Penhune, 2013)和皮层下(脑干) (Bidelman, Weiss, Moreno, & Alain, 2014; Musacchia, Strait, & Kraus, 2008; Wong, Skoe, Russo, Dees, & Kraus, 2007)结构和功能的变化, 使得音乐训练者从低级的听觉感知到高级的多层级加工(注意、记忆等)能力得到增强(Moreno & Bidelman, 2014), 从而促进语言加工。然而, 音乐训练者在加工语言的韵律信息, 尤其是在加工堪称中国语言韵律典范的近体诗时, 是否也具有优势尚未可知。本研究拟使用七言绝句考察此问题。

韵律在语言中包括语速、语调、声调(声调语言独有)、重读等超音段特征(Hausen, Torppa, Salmela, Vainio, & Särkämö, 2013), 与音强、时长和基频等声学参数有关。在口语句子理解中, 韵律线索可以用来建立句法结构并提供语义信息, 是语音习得和感知不可缺少的元素(Dahan, 2015; Julia et al., 2013; Kotz et al., 2003; Steinhauer, Alter, & Friederici, 1999)。音乐同语言一样, 作为人类获取听觉信息的主要来源, 可以通过音调的高低、节奏的快慢等特征实现表情、表意、表态的功能(Dilley, Mattys, & Vinke, 2010; Juslin & Laukka, 2003; Patel, 2003; 蒋存梅, 2016)。有研究发现作曲家对音乐旋律的创作受言语韵律的影响(Patel, Iversen, & Rosenberg, 2006), 具有音乐障碍的先天性失乐症患者(尤其是音高加工障碍)在感知、识别言语韵律时更加困难(Liu, Patel, Fourcin, & Stewart, 2010; Nan, Sun, & Peretz, 2010), 表明音乐和语言在韵律方面有着诸多联系(Hausen et al., 2013; Patel, Peretz, Tramo, & Labreque, 1998; Patel & Daniele, 2003)。

此外, 音乐家加工与韵律相关的声学线索时具有优势(南云, 2017)。例如, 音乐家在感知音高和声调变化时的脑干编码更精确(Bidelman, Gandour, & Krishnan, 2011; Lee & Hung, 2008), 并且在皮层电反应上对音高以及音节水平语音时间结构(如音长、声音开始时间)的细微变化更加敏感(Kühnis,

收稿日期: 2019-12-04
* 辽宁省教育厅自然科学基金(LJ2019016); 辽宁省自然科学基金(20180550313); 中国科学院行为科学重点实验室开放课题基金(Y5CX052003)资助。
通信作者: 李卫君, E-mail: li_wj@126.com

Elmer, Meyer, & Jäncke, 2013; Partanen, Vainio, Kujala, & Huotilainen, 2011; Wu et al., 2015)，其中纵向研究结果表明音乐能力在语言领域的迁移效应更可能源于训练而非基因倾向(Chobert, François, Velay, & Besson, 2012; Nan et al., 2018)。长期的音乐训练不仅提高了对低水平声学信息加工的敏感性，而且对一些高级认知加工起促进作用，如注意(Wang, Ossher, & Reuter-Lorenz, 2015)、言语记忆(Franklin et al., 2008)、执行功能(陈杰，刘雷，王蓉，沈海洲，2017)、语音分段(François, Chobert, Besson, & Schön, 2012)、语言表达技能(Milovanov, Huotilainen, Välimäki, Esquef, & Tervaniemi, 2008)等。基于以往结果，OPERA 假说解释了为何音乐训练可能通过共享的神经资源对语言加工起促进作用：当音乐对这两个领域共享的感官和认知加工机制提出的要求比语音更高时，音乐能力可以增强语音加工；这些更高要求的结合以及与音乐能力相关的情感奖励，重复和注意激活了神经可塑性，改变了影响语音加工的神经结构和功能，从而促进语言加工(Patel, 2014)。因此，长时间系统的音乐训练可能提高人们对言语韵律的敏感性。

诗歌，作为音乐和语言的中介(interlude)，其包含的押韵、平仄(通过声调体现)等格律要素也属于韵律信息。诗歌韵律往往体现了其民族语言的特色，是通过不断地探索和尝试，最终确定下来的几近完美的格式和规律，是语言韵律的结晶。中国的近体诗因为有着严格且成熟的格律要求，其传达的语言韵律之美更是达到了世界诗歌史的巅峰(吴洁敏，朱宏达，2001)。在七言绝句这一近体诗中，每篇为四句两联，每句要求“平仄”交替配置，一联的上下句中平仄也要相互对立，每一联为一个周期。此外，七言绝句通常第一句、第二句、第四句最后一个字的韵母相同，且声调为平声(即押韵)。这些格律要素中变化的规律性和整体性构成了语流上的和谐韵律，增强了诗句的音乐性、表现力和共鸣感。因此某些位置上违反格律要求的变化(如，变化韵脚的声调和韵母)会破坏诗句的韵律美。那么，音乐训练是否会影响听者感知并加工由声调和韵母变化引起的诗句整体的韵律违反？这是我们关注的第一个问题。

句末是句法、语义、韵律等信息完整呈现的位置，听者/读者在此位置的加工负荷相比句内更大，并存在更高水平的整合加工(Just & Carpenter, 1980; Kuperberg, Kreher, Goff, McGuire, & David, 2006)。已有行为和眼动研究发现，句末相较于其他位置的阅读时间更长(de Vincenzi et al., 2003; Just, Carpenter, & Woolley, 1982)、回视和眼跳更多(Camblin, Gordon, & Swaab, 2007; Rayner, Kambe, & Duffy, 2000)。采用 ERP 技术探究绝句韵律边界加工的研究发现，与诗句内部各韵律边界诱发的脑电效应(CPS)不同，在诗句末边界诱发了波幅更大的正波(P3)，可能反映了句法终止和完整语言单元的结束(李卫君，杨玉芳，2010)。值得注意的是，一系列研究将句法、语义信息违反设置在句内，却在句末发现违反相比合适条件诱发了波幅更大且更持久的中后部分布的 N400 效应(Bohan, Leuthold, Hijikata, & Sanford, 2012; Ditman, Holcomb, & Kuperberg, 2007; Hagoort, 2003; 金花等, 2009; Molinaro, Vespignani, & Job, 2008; Osterhout, & Holcomb, 1993; Osterhout & Mobley, 1995; Osterhout & Nicol, 1999)。更大波幅的 N400 可能反映了被试在句末完成句法、语义整合时，整合难度更大。也就是说，被试在将前面出现的违反信息整合进整个语境时耗费了更多的认知资源。那么，对于句子内部出现韵律违反，在句末位置是否也会探测到像句法、语义违反一样的整合加工过程，并诱发相似的脑电效应？这是本研究探讨的另一个问题。

综上，与句中不同，句末位置存在基于整个句子的整合加工过程。以往研究主要考察了句法和语义的句末整合加工，尚未有研究考察韵律信息的句末整合机制。目前采用违反范式探究言语韵律的研究主要考察了声调、韵母信息违反的即时加工过程。大多研究采用行为和脑电技术，使用不同的材料(如，单字对、成语、高限制语义句子、诗句等)和不同的实验任务(如，词汇判断、语义判断、词汇决策等)，探究声调和韵母在词汇识别和限制语义整合加工过程中的时间进程(Hu, Gao, Ma, & Yao, 2012; Huang, Liu, Yang, Zhao, & Zhou, 2018; Li, Wang, & Yang, 2014; Schirmer, Tang, Penney, Gunter, & Chen, 2005)。虽然尚未得出一致结论，但是大多数研究发现相较于声调信息，韵母信息的加工相对更快；即使母语是声调语言者，韵母在词汇识别和语义限制方面仍具有更强的影响力和约束力(Tong, Francis, & Gandour, 2008; Hu et al., 2012; Huang et al., 2018; Li et al., 2014)。韵母信息的即时加工优势在句末整合时是否仍然会有体现？我们拟在本研究探讨。同时，长期高强度的音乐训练使音乐训练者的生理基础及功能发生了改变。以往研究者通过变化句末单

词的音高或音节时长调节句子韵律，观察音乐训练者加工韵律信息的时间进程，发现音乐训练者对韵律变化的反应更加灵敏，加工过程更加高效(Magne, Schön, & Besson, 2006; Marques, Moreno, Luís Castro, & Besson, 2007; Marie, Magne, & Besson, 2011; Schön, Besson, & Magne, 2004; Zioga, Luft, & Bhattacharya, 2016)。但是，目前还不清楚音乐训练对诗句内发生韵律违反，诗句末完成韵律信息整合时是否也会起到促进作用。基于此，本研究拟采用七言绝句为材料，利用押韵判断任务，考察音乐训练组和对照组在诗句末整合韵律信息的认知过程。为了避免诗句末整合与操纵的刺激本身诱发效应的混淆，本研究通过改变绝句第一联末字的声调和韵母造成押韵违反，在绝句末考察两组被试完成绝句韵律信息整合引起的脑电效应。我们预期诗句内的声调和韵母违反均可能引起句末整合困难：相比于合适条件，诗句内的声调和韵母违反会在句末引发反映整合困难的 N400 效应。此外，音乐训练组由于其长期音乐训练导致对韵律信息更为敏感，会相比对照组诱发更早、波幅更大的脑电效应。

2　方法

2.1　被试

从高校招募受过和未受过音乐训练的普通话母语者各 25 名(9 男，16 女，年龄范围均为 18~26 岁，其中音乐训练组年龄为 20.56 ± 1.90 岁，对照组年龄为 20.84 ± 1.85 岁)。两组被试均为右利手，无精神病史，视力或矫正视力及听力正常。音乐训练组均自我报告从 7 岁或以前开始接受音乐训练，有十年以上专业音乐训练经历，且近期平均每天至少有 1 小时的音乐练习。两组被试在年龄、性别、教育程度上进行匹配(如表 1 所示)。正式实验开始前，每位被试完成瑞文高级智力测验简化版(Arthur & Day, 1994)。统计发现，音乐训练者(M = 7.52, SD = 1.90)和普通人(M = 8.44, SD = 1.87)流体智力水平无显著差异，t (48) = 1.73, p = 0.091(双侧)。研究得到辽宁师范大学伦理委员会批准，且被试在实验前签署了知情同意书，在实验后获得一定报酬。

表 1　被试人口学资料

编号	性别		年龄(岁)		受教育程度		训练开始年龄(岁)		训练时长(年)		乐器	
	M	N	M	N	M	N	M	N	M	N	M	N
1	男	男	18	18	本科生	本科生	7	无	11	0	竹笛	无
2	男	男	18	19	本科生	本科生	7	无	11	0	爵士鼓、小军鼓	无
3	男	男	19	19	本科生	本科生	7	无	12	0	钢琴	无
4	男	男	19	20	本科生	本科生	7	无	11	0	钢琴	无
5	男	男	19	20	本科生	本科生	7	无	10	0	二胡、双排键	无
6	男	男	21	20	本科生	本科生	7	无	10	0	竹笛、葫芦丝	无
7	男	男	21	21	本科生	本科生	6	无	10	0	二胡、钢琴	无
8	男	男	22	22	本科生	本科生	7	无	15	0	二胡	无
9	男	男	23	23	本科生	本科生	7	无	15	0	钢琴、吉他	无
10	女	女	19	19	本科生	本科生	7	无	11	0	扬琴	无
11	女	女	19	19	本科生	本科生	6	无	12	0	扬琴	无
12	女	女	19	19	本科生	本科生	6	无	13	0	扬琴	无
13	女	女	19	20	本科生	本科生	6	无	13	0	琵琶	无
14	女	女	19	20	本科生	本科生	5	无	10	0	琵琶	无
15	女	女	20	20	本科生	本科生	7	无	13	0	钢琴、古筝	无
16	女	女	20	20	本科生	本科生	7	无	13	0	钢琴、吉他	无
17	女	女	20	20	本科生	本科生	6	无	14	0	钢琴	无
18	女	女	21	21	本科生	本科生	7	无	14	0	钢琴	无
19	女	女	21	21	本科生	本科生	5	无	16	0	钢琴、大提琴	无
20	女	女	21	22	本科生	本科生	4	无	17	0	钢琴、古筝	无
21	女	女	21	22	本科生	本科生	4	无	17	0	钢琴、小号、古筝	无
22	女	女	23	23	本科生	本科生	7	无	16	0	钢琴、古筝、架子鼓	无
23	女	女	23	23	本科生	本科生	7	无	16	0	钢琴、双排键、琵琶	无
24	女	女	23	24	研究生	研究生	5	无	18	0	钢琴、双排键、琵琶	无
25	女	女	26	26	研究生	研究生	7	无	12	0	钢琴、葫芦丝、长笛	无

注：M = 音乐训练组(musician), N = 对照组(nonmusician)

2.2 实验材料

实验材料采用以往研究使用的160首低熟悉度的七言绝句(Li et al., 2014)。由于韵母完全相同且熟悉度低的诗句数量有限，因此我们使用的绝句除了第一句、第二句、第四句最后一个字的主要韵母相同的材料外，还包括了三种其他押韵方式的材料：eng, ing：如“生，轻，行”; en, in：如“侵，阴，深”; ou, iu：如“流，幽，鸥” (占总材料 20.6%)。低熟悉度量化标准具体为：首先通过网络和书籍选择实验者不熟悉的 300 首七言绝句，然后进行预实验。预实验中，16 名大学生阅读依次呈现的绝句，并要求他们判断这些绝句是否熟悉，并在 7 点量表上表明该绝句意义上的可理解程度。最终我们选择被试认为不熟悉并且意义性较低(5 分及其以下，主要为了平衡最小语义干扰和开展 ERP 实验所需试次)的 240 首七言绝句作为正式实验材料。通过改变诗句第一联最后一个汉字的声调或韵母构成4种实验条件：声调、韵母均合适(V+T+)，声调合适、韵母违反(V−T+)，声调违反、韵母合适(V+T−)，声调违反、韵母违反(V−T−)。具体为，声调违反条件统一由 4 声代替原诗句的 1 声或 2 声；韵母违反通过变化其韵母部分，使其与第一联第一句中最后一个字的韵母不同。在我们操纵韵母违反时并没有细致地从介音、主要元音和韵尾的角度考虑，仅是以违反押韵规则为标准将韵母看作一个整体进行变化，使其与符合押韵规则的韵母在感知上存在显著差别(如图 1 所示)。双违反条件则同时将其声调变为 4 声，并使韵母发生如上所述变化。实验最终形成 640 首诗句。为避免被试对特定位置产生预期，实验还包括其余 80 首七言绝句作为填充材料：声调、韵母均合适(40 首)，以及通过变化第二联最后一个汉字形成的声调合适、韵母违反(13 首)，声调违反、韵母合适(13 首)，声调、韵母均违反(14 首)。采用拉丁方将实验材料进行分组，使得同一诗句的不同实验条件分别出现在四个列表(list)中；将填充材料加入后对同一列表中诗句进行伪随机排序。所有材料由一名男性发音人正常朗读，采样率为 22 kHz。对于声调和韵母全合适的绝句，发音人会完整朗读每首绝句；声调和韵母任何一个维度违反的绝句，发音人仅朗读诗句第一联，然后通过语音拼接技术将第一联末字拼接到完全合适的绝句中。因此，对于同一绝句的不同条件，仅第一联末字不同，其他部分完全相同。通过统计发现，该发音人朗读每首合适绝句总体发音时间在 14 s 到 17 s 之间。

2.3 实验设计及程序

采用 2 被试类型(音乐训练组、对照组) × 2 声调(合适、违反) × 2 韵母(合适、违反)三因素混合实验设计。被试类型为组间变量，声调、韵母为组内变量。

实验在光线柔和，安静舒适的房间内进行。被试坐于液晶显示屏(规格为23″，刷新率为 60 Hz)前，佩戴 Panasonic RP-HS47 挂耳式耳机完成实验。实验共有 6 个区组(block)，每个区组包括 40 个试次，同一条件的实验材料连续呈现不超过三次。每个被试听 240 首诗(160 首实验材料, 80 首填充材料)。每个区组大约用时 10 分钟，区组的间歇被试根据自身情况休息。

被试经练习熟练掌握实验要求后，开始正式实验。练习前明确告知被试押韵规则。在练习中，我们将 eng, ing; en, in; ou, iu 三种类型的押韵材料增加在内，并告知被试受历史影响这些诗句仍属于押韵合适。每个试次中，首先同时呈现注视点(“＋”)和提示音 300 ms，随后“＋”继续呈现在屏幕中央以减少被试眼动，同时播放一首七言绝句，声音刺激播放完毕后会在 1/3 的试次后出现探测界面，要求被试既快又准地判断诗句是否押韵。押韵按“F”键，不押韵按“J”键，并在被试间进行按键平衡。试次间的随机间隔为 1000~1200 ms (如图 1 所示)。

2.4 数据采集和分析

采用 ANT 设备(ANT Neuro)，按照国际 10-20 系统扩展的 64 导电极帽记录 EEG。信号记录的采样率为 500 Hz，以 CPz 为在线参考。电极 M1 和 M2 分别置于左侧和右侧乳突。电极与头皮之间的阻抗小于 5 kΩ，在线记录的滤波带通为 0.01~100 Hz。离线分析时从各导联的脑电数据中减去双侧乳突的平均数作再参考。采用 Brain Vision Analyzer 2.0 软件进行数据处理，对脑电信号进行滤波(高通为 0.01 Hz，低通为 30 Hz)。脑电分析锁定的起始点为每首诗中第二联最后一个字的开始位置，截取此关键位置之前 200 ms 和之后 1000 ms 的数据存为 EEG 数据，并进行基线矫正。排除眼动伪迹及电位超过 ± 80 μV 的其它伪迹后，每种条件下的有效试次均在 30 次以上。

基于诱发脑电效应的分布情况及以往相关研究(Hagoort, 2003；李卫君，刘梦，张政华，邓娜丽，邢钰珊，2018)。对中线和左右两半球 100~300 ms 和 300~750 ms 两个时间窗口进行被试类型(音乐训练组、对照组) × 韵母(合适、违反) × 声调(合适、

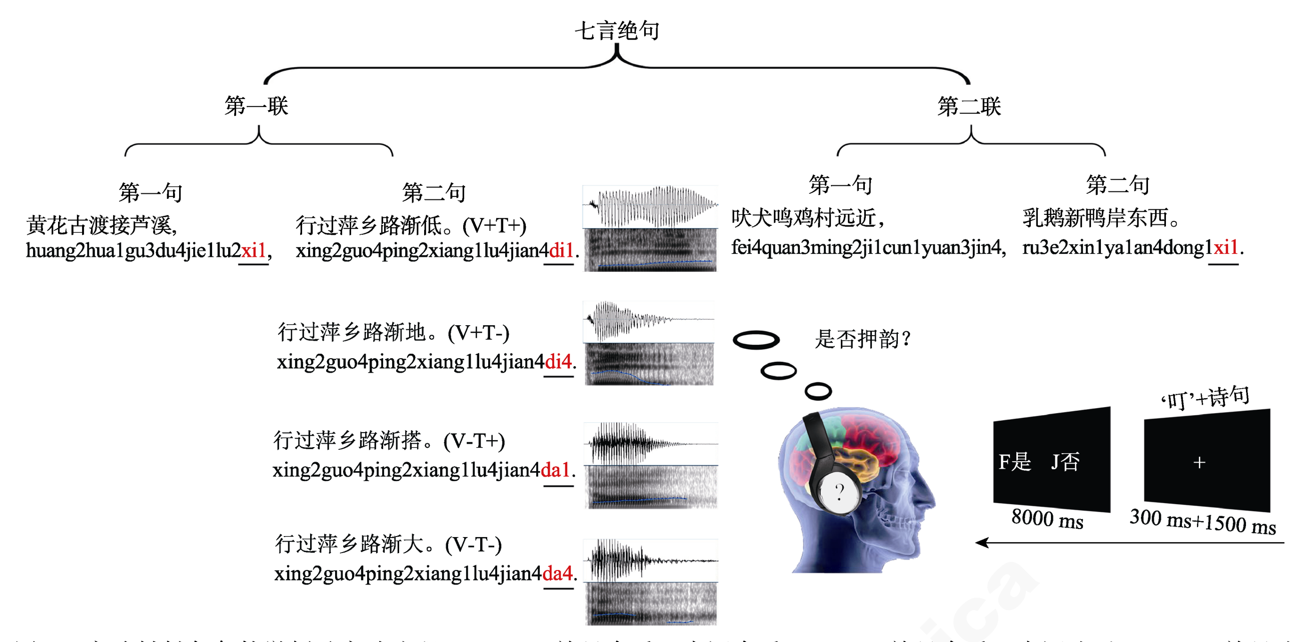


图 1　实验材料各条件举例及实验流程。V+T+：韵母合适、声调合适; V+T−：韵母合适、声调违反; V−T+：韵母违反、声调合适; V−T−：韵母违反、声调违反

违反)重复测量方差分析。中线上 3 个兴趣区包括额区(Fz, FCz), 中央区(Cz, CPz)和顶区(Pz, POz)。两侧增加半球(左, 右)因素, 左右半球共划分 6 个兴趣区, 包括左前(F1, F3, F5, FC1, FC3, FC5), 右前(F2, F4, F6, FC2, FC4, FC6), 左中(C1, C3, C5, CP1, CP3, CP5), 右中(C2, C4, C6, CP2, CP4, CP6), 左后(P1, P3, P5, PO3, PO5, O1)和右后(P2, P4, P6, PO4, PO6, O2)。用 Greenhouse-Geisser 对不符合球形检验的 P 值进行校正(Greenhouse and Geisser, 1959)。

3　结果

3.1　行为结果

对正确率进行独立样本 t 检验发现, 音乐训练组(M = 83.24%, SE = 6.64%)和对照组(M = 82.64%, SE = 9.15%)完成押韵合适性判断的正确率不存在显著差异, $t(48) = 0.27$, $p = 0.791$。进行 2 (被试类型：音乐训练组、对照组) × 2 (韵母：合适、违反) × 2 (声调：合适、违反)重复测量方差分析发现, 韵母违反条件下的正确率(M = 90.70%, SE = 1.90%)高于韵母合适条件(M = 79.90%, SE = 1.60%), $F(1, 48) = 26.09$, $p < 0.001$, $\eta_p^2 = 0.35$, 声调违反条件下的正确率(M = 87.10%, SE = 2.00%)高于声调合适条件(M = 83.60%, SE = 1.20%), $F(1, 48) = 3.93$, $p = 0.053$, $\eta_p^2 = 0.08$(边缘显著) (如图 2 所示)。

对反应时进行 2 (被试类型：音乐训练组、对照组) × 2 (韵母：合适、违反) × 2(声调：合适、违反)重复测量方差分析发现, 韵母违反条件下的反应时(M = 838.79, SE = 51.66)短于韵母合适条件(M = 933.44, SE = 45.06), $F(1, 48) = 5.27$, $p = 0.026$, $\eta_p^2 = 0.10$。声调违反条件下的反应时(M = 803.91, SE = 46.75)短于声调合适条件(M = 968.32, SE = 53.28), $F(1, 48) = 11.50$, $p = 0.001$, $\eta_p^2 = 0.19$。被试和声调的交互作用显著, $F(1, 48) = 5.85$, $p = 0.019$, $\eta_p^2 = 0.11$。简单效应分析表明, 仅在音乐训练组中发现声调违反条件下的反应时显著短于声调合适条件, $F(1, 48) = 16.88$, $p < 0.001$, $\eta_p^2 = 0.26$ (如图 2 所示)。

3.2　ERP 结果

3.2.1　100~300 ms 时间窗

统计分析发现, 中线上, 被试、韵母和声调交互作用显著, $F(1, 48) = 7.59$, $p = 0.008$, $\eta_p^2 = 0.14$。简单效应分析表明, 音乐训练组在声调合适时, 韵母违反相比韵母合适诱发了波幅更大的正效应, $F(1, 48) = 9.84$, $p = 0.003$, $\eta_p^2 = 0.17$; 韵母合适时, 声调违反相比声调合适诱发了波幅更大的正效应(边缘显著), $F(1, 48) = 3.80$, $p = 0.057$, $\eta_p^2 = 0.07$; 韵母违反时, 声调违反相比声调合适诱发了波幅更小的正效应, $F(1, 48) = 10.68$, $p = 0.002$, $\eta_p^2 = 0.18$。对照组在以上所有实验条件下均未诱发明显正效应(ps > 0.1)。

在两侧, 韵母违反(M = –0.97, SE = 0.11)相比韵母合适(M = –1.22, SE = 0.15)诱发了波幅更大的正效应, $F(1, 48) = 4.04$, $p = 0.05$, $\eta_p^2 = 0.08$。被试、韵母和声调交互作用显著, $F(1, 48) = 7.54$, $p = 0.008$,

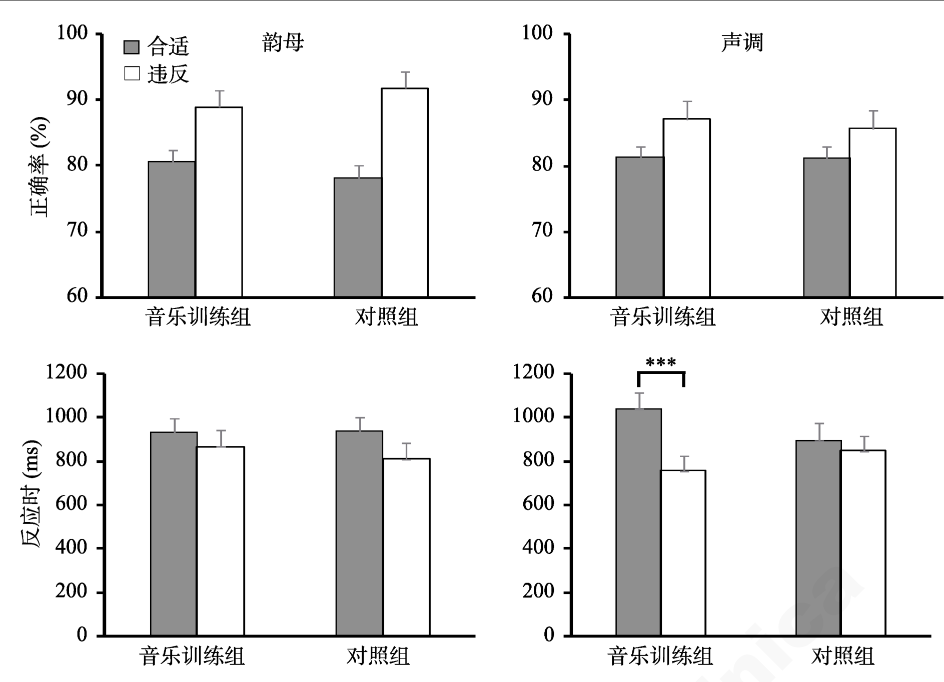


图 2 音乐训练者(25 人)和相应对照组(25 人)完成押韵判断任务时，韵母(左)和声调(右)在合适和违反条件下的平均正确率(上)和反应时(下)。

η_p^2 = 0.14。简单效应分析表明，音乐训练组在声调合适时，韵母违反相比韵母合适诱发了波幅更大的正效应，$F(1, 48) = 15.41, p < 0.001, \eta_p^2 = 0.24$；韵母合适时，声调违反相比声调合适诱发了波幅更大的正效应，$F(1, 48) = 7.27, p = 0.010, \eta_p^2 = 0.13$；韵母违反时，声调违反相比声调合适诱发了波幅更小的正效应，$F(1, 48) = 7.37, p = 0.009, \eta_p^2 = 0.13$。对照组在以上所有实验条件下均未诱发明显正效应($ps > 0.1$)。被试、韵母、半球和脑区的交互作用显著，$F(2, 96) = 3.91, p = 0.033, \eta_p^2 = 0.08$。简单效应分析表明，左半球中部，音乐训练组韵母违反相比韵母合适诱发了明显正效应，$F(1, 48) = 4.69, p = 0.35, \eta_p^2 = 0.09$；对照组韵母违反和韵母合适无显著差异($F < 1$) (如图 3, 4 所示)。

总之，该时间窗口的结果表明，仅音乐训练组在绝句末会对绝句内部声调和韵母的一致性进行深入分析和整合。

3.2.2 300~750 ms 时间窗

统计分析发现，中线上，韵母与脑区交互作用显著，$F(2, 96) = 9.96, p = 0.001, \eta_p^2 = 0.17$。简单效应分析表明，韵母违反相比于韵母合适在中部[$F(1,48) = 16.45, p < 0.001, \eta_p^2 = 0.26$]和后部[$F(1,48) = 24.87, p < 0.001, \eta_p^2 = 0.34$]诱发了波幅更大的负效应。声调与脑区交互作用显著，$F(2, 96) = 5.61, p = 0.008, \eta_p^2 = 0.11$。简单效应分析表明，声调违反相比于声调合适在中部[$F(1,48) = 16.10, p < 0.001, \eta_p^2 = 0.25$]和后部[$F(1,48) = 45.65, p < 0.001, \eta_p^2 = 0.49$]诱发了更大的负效应。被试、韵母和声调交互作用显著，$F(1, 48) = 6.66, p = 0.013, \eta_p^2 = 0.12$。简单效应分析表明，音乐训练组在声调违反时，韵母违反相比韵母合适诱发了波幅更大的负效应，$F(1, 48) = 5.14, p = 0.028, \eta_p^2 = 0.10$；韵母违反时，声调违反相比声调合适诱发了波幅更大的负效应，$F(1, 48) = 12.94, p = 0.001, \eta_p^2 = 0.21$。对照组在声调合适时，韵母违反相比韵母合适诱发了波幅更大的负效应，$F(1, 48) = 15.07, p < 0.001, \eta_p^2 = 0.24$；韵母合适时，声调违反相比声调合适诱发了波幅更大的负效应，$F(1, 48) = 10.88, p = 0.002, \eta_p^2 = 0.19$。

在两侧，韵母与脑区交互作用显著，$F(2, 96) = 10.97, p = 0.001, \eta_p^2 = 0.19$。简单效应分析表明，韵母违反相比于韵母合适在中部[$F(1,48) = 13.38, p = 0.001, \eta_p^2 = 0.22$]和后部[$F(1,48) = 22.22, p < 0.001, \eta_p^2 = 0.32$]诱发了波幅更大的负效应。声调与脑区交互作用显著，$F(2, 96) = 12.54, p < 0.001, \eta_p^2 = 0.21$。简单效应分析表明，声调违反相比于声调合适在中部[$F(1,48) = 21.35, p < 0.001, \eta_p^2 = 0.31$]和后部[$F(1,48) = 27.69, p < 0.001, \eta_p^2 = 0.37$]诱发了波幅更大的负效应。被试、韵母和声调交互作用显著，$F(1, 48) = 5.60, p = 0.022, \eta_p^2 = 0.10$。简单效应分析表明，音乐训练组在声调违反时，韵母违反相比韵母合适诱发了波幅更大的负效应(边缘显著)，$F(1, 48) = 2.92, p = 0.094, \eta_p^2 = 0.06$；韵母违反时，声调违反

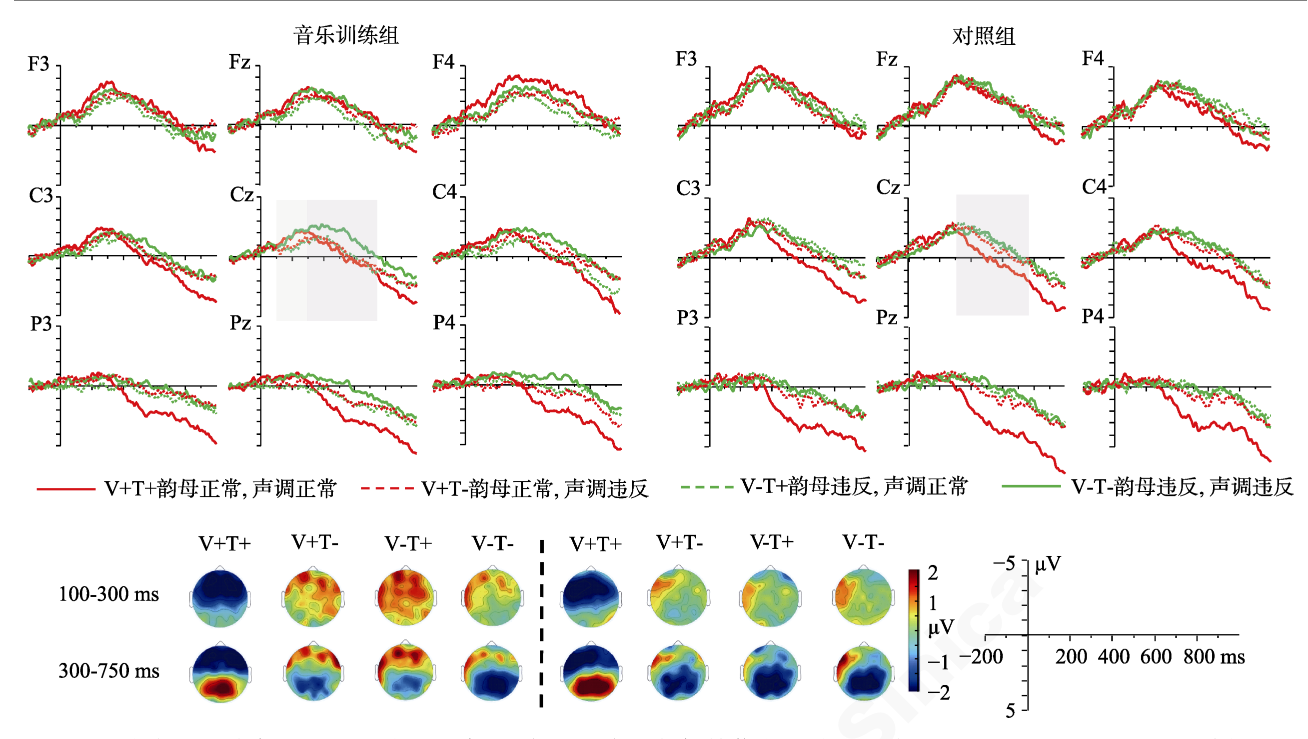


图 3　25 个音乐训练者(左)和 25 个对照者(右)加工七言绝句押韵信息的总平均波形(上)和地形图(下)。ERP 分析的开始时间为第二联末字开始位置。地形图为全合适条件的地形图以及各违反条件与全合适条件的差异波地形图。

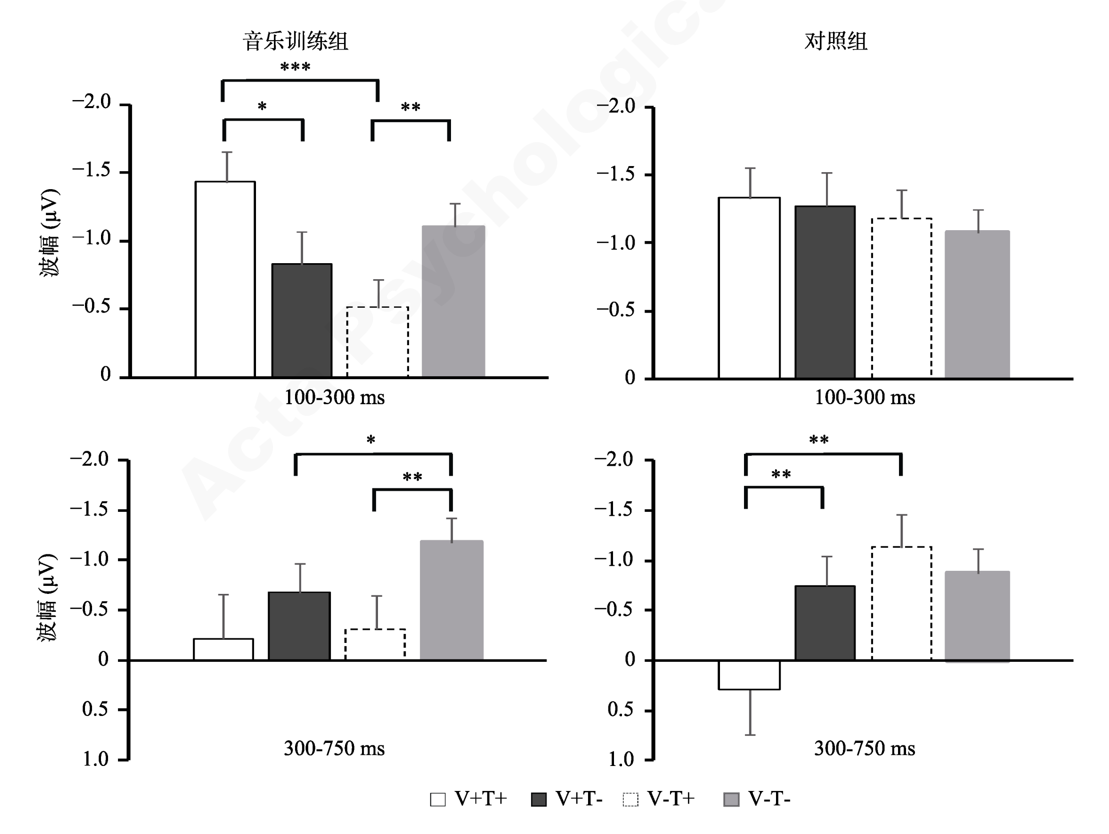


图 4　25 个音乐训练者(左)和 25 个对照组被试(右)加工七言绝句押韵信息时大脑两侧 100~300 ms (上)和 300~750 ms (下)诱发脑电效应的平均波幅。V+T+：韵母正常，声调正常；V+T−：韵母正常，声调违反；V−T+：韵母违反，声调正常；V−T−：韵母违反，声调违反。

相比声调合适诱发了波幅更大的负效应, $F(1, 48) = 9.65$, $p = 0.003$, $\eta_p^2 = 0.17$。对照组在声调合适时，韵母违反相比韵母合适诱发了波幅更大的负效应, $F(1, 48) = 12.04$, $p = 0.001$, $\eta_p^2 = 0.20$；韵母合适时，

声调违反相比声调合适诱发了波幅更大的负效应，$F(1, 48) = 8.27$, $p = 0.006$, $\eta_p^2 = 0.15$ (如图 3 所示)。

总之，在此时间窗口两类被试都会在诗句末对诗句内发生的声调和韵母违反完成进一步分析，不过音乐训练组在声调和韵母违反时才对韵母和声调的一致性进行分析，而对照组则在声调和韵母合适时才会深入分析韵母和声调的一致性。

4 讨论

本研究通过给被试呈现不熟悉的七言绝句，利用押韵判断任务，考察音乐训练组和对照组在绝句末完成韵律信息的整合过程。行为结果表明，音乐训练组和对照组在完成任务的正确率上未出现显著差异，且相对声调和韵母合适条件，二者加工声调和韵母违反时错误率更低；仅在音乐训练组中发现声调违反条件下的反应时显著短于声调合适条件。ERP 结果发现，音乐训练组和对照组完成诗句韵律信息的整合加工存在明显差异。首先，在 100~300 ms，对照组没有诱发任何显著效应，音乐训练组在声调/韵母合适条件下，对韵母/声调的一致性进行深入分析，并诱发了更大的正波。在韵母违反条件下，声调违反相比声调合适诱发了更小的正波。此外，在 300~750 ms，两组被试均在绝句末对绝句内出现的韵母和声调违反进行整合分析。不过，对照组在声调/韵母合适条件下，对韵母/声调的一致性进行深入分析并诱发负波，而音乐训练组在声调/韵母违反条件下，对韵母/声调的一致性进行分析并诱发负波。以上结果表明，两组被试均会在绝句末进一步对诗句内部押韵违反进行深入分析，不过音乐训练影响了该分析的具体过程，由此说明长期的音乐训练可能改变了大脑皮层对韵律加工的反应模式，促使音乐训练者对语音韵律的加工更快速，并且对声调信息的变化更加敏感。

4.1 诗句韵律信息的诗句末整合

行为结果发现，合适条件比违反条件的正确率低、反应时慢的结果与以往研究一致(Schirmer et al., 2005，使用高完型概率的粤语句子；Hu et al., 2012，使用成语；及 Li et al., 2014，使用诗句)。当前实验对于合适条件和违反条件在正确率和反应时上产生的差异，可能是因为，被试只要探测到第一联末出现不合适情况，就能够判定该绝句属于押韵违反，相对容易；但是对于合适情况则需要听完整个绝句，确定绝句每一个地方都合适，才能做出合适判断，相对更难。

更重要的是，不论是音乐训练组还是对照组，均在诗句末稳定探测到了由绝句内部押韵违反引起的整合加工过程。由此表明，人们不仅在句末会完成对由句中句法、语义违反引起的整合加工(Bohan et al., 2012; Ditman et al., 2007; Molinaro et al., 2008)，也会对韵律违反进行相似的整合过程。具体表现在绝句末字开始后大约 300 ms，诗句内声调和韵母违反相比合适条件均诱发了中后部分布的负波，且这一效应一直持续到 750 ms 或更久，与以往探究句法、语义句末整合加工的研究结果一致(Bohan et al., 2012; Molinaro et al., 2008; Osterhout & Holcomb, 1993; Osterhout & Nicol, 1999)。当前研究中负波存在三种可能的解释：第一，其潜伏期、脑区分布与经典 N400 一致，多解释为语义加工(Kutas & Federmeier, 2011)。先前研究发现即使要求被试仅关注实验材料的韵律信息(音长或押韵)，语义违反仍会诱发 N400，体现了听者对语音材料语义加工的自动化(Magne et al., 2007; Marie et al., 2011; Perrin & García-Larrea, 2003)。本研究中虽选取熟悉度低的诗句为材料，但在汉语中声调和韵母均对汉字的识别起重要作用。因此诗句中出现的声调、韵母信息引起的韵律违反导致了诗句末语义信息整合困难。第二，该负波直接反映了对韵律信息的整合困难。有研究发现当要求被试对依次呈现的单词进行押韵判断时，即使是毫无意义的假词，不押韵的词也会诱发后部分布的 N400 (Coch, Grossi, Skendzel, & Neville, 2005; Praamstra, & Stegeman, 1993)。这表明 N400 不仅反映语义整合，也对词汇加工过程中的音韵信息非常敏感(Chen et al., 2016; Li et al., 2014)。与此解释一致，当前结果可能表明在押韵判断任务下，听者会在诗句末再次审视整首诗句的韵律信息，当绝句内部出现与预期韵律不符的刺激时，会在诗句末增加对整首诗韵律信息的整合难度。第三，针对句中不同的信息违反(句法、语义、韵律)，在句末仍会稳定的诱发在潜伏期、分布位置等无明显差异的负波(Bohan et al., 2012; Molinaro et al., 2008; Osterhout & Holcomb, 1993; Osterhout & Nicol, 1999)。说明在任务导向下，在句末位置会对这些违反信息进行无差别的再整合，此负波可能反映的是将有错误的句子元素整合进连贯整体的语音表征中有困难，更大波幅的负波表明整合难度更大。不论诗句末负波解释为何种过程，本研究都表明，两类被试会在诗句末对诗句内出现的声调、韵母违反信息进行整合加工。

以往采用违反范式探究言语韵律的研究主要考察了声调、韵母信息违反的即时加工过程。大多数研究发现相较于声调信息，韵母信息的加工相对更快；即使母语是声调语言者，韵母在词汇识别和语义限制方面仍具有更强的影响力和约束力(Tong, Francis, & Gandour, 2008; Hu et al., 2012; Huang et al., 2018; Li et al., 2014)。与本研究最为相关的研究(Li et al., 2014，使用普通被试)也发现了韵母信息的加工优势。研究者通过变化七言绝句第一联末字声调和韵母信息产生韵律违反，发现在 300~500 ms 韵母违反相比韵母合适诱发了波幅更大的负波(N450)；在 600~1000 ms，声调违反相比声调合适在中后部诱发波幅更大的正波(LPC)，并且在声调合适条件下，韵母违反相比韵母合适在后部诱发波幅更大的正波(LPC)。本研究主要考察了音乐训练组和对照组被试诗句末的韵律整合过程。对照组的结果显示，在 300~750 ms，声调合适条件下，韵母违反相比韵母合适诱发波幅更大的负波；韵母合适条件下，声调违反相比声调合适诱发波幅更大的负波。两研究结果比较，可以发现在对诗句中出现的押韵违反信息的即时加工过程中，人们探测韵母违反相较于声调违反更加快速，声调和韵母的交互作用发生在晚期时间窗口(600~1000 ms)。诗句末韵律整合加工是对绝句内出现的声调、韵母信息违反进行的深入整合，在绝句末关键字出现后 300 ms 即同时开始了声调和韵母的整合加工，并诱发了中后部分布的持久负波。总之，对照组诗句末韵律的整合过程不同于诗句内韵律信息的即时加工，体现了诗句末韵律加工的独特过程。

4.2　音乐训练组的加工优势

本研究发现，音乐训练组和对照组均在声调/韵母合适的情况下，对韵母/声调的一致性进行深入分析，只是音乐训练组完成这一过程的时间(100~300 ms)相比对照组(300~750 ms)更早，且主要体现为一个类似 P200 的早期正波。以往研究表明，长期高要求的音乐训练会引起大脑环路结构和功能的变化，从而通过低水平的感知和高水平认知控制的一系列增强来实现语言加工优势效应，使得音乐训练者比非音乐训练者在一些任务中表现更好(Bidelman, et al., 2014; James et al., 2014; Nan et al., 2018; Zioga et al., 2016)。本研究中音乐训练组在 100~300 ms 时间窗诱发的正效应，与以往研究中音乐训练者在音乐或语言领域加工优势的发现一致。例如，有研究发现与非音乐家相比，音乐家加工句子韵律(音节时长)时在 150~250 ms 时间窗诱发了更大的正效应(Marie et al., 2011)。不仅是长期的音乐训练，有研究发现短期的声音感知训练也会导致早期时间窗正波波幅的增加(Atienza, Cantero, & Dominguez-Marin, 2002; Reinke, He, Wang, & Alain, 2003; Tremblay, Kraus, McGee, Ponton, & Otis, 2001)，表明此时间段的正波可能体现了音乐训练的神经可塑性(Marie et al., 2011; Reinke et al., 2003; Shahin, Bosnyak, Trainor, & Roberts, 2003)，更大波幅的正波可能反映了神经同步性的增强(Shahin et al., 2003)或是更多神经元的集合反应(Reinke et al., 2003)。以往研究者将此时间段的正波定义为 P200，认为其属于独立于注意状态的外源成分，反映了对低水平声学刺激特征的自动加工(Crowley & Colrain, 2004; Shahin, Roberts, Pantev, Trainor, &Ross, 2005)，快速注意捕获过程(Carpenter, Cranford, Hymel, de Chicchis, & Holbert, 2002; Fan et al., 2016)，也可能与早期音韵加工负荷有关(Huang, Yang, Zhang, & Guo, 2014; Marie et al., 2011)。这些解释都是在对目标刺激的即时加工中得出的。本研究中绝句末字不存在条件间差异，虽然诱发的正波在时间上出现的较早，但是实验控制的条件发生在绝句内，所以它更可能反映的是对诗句内发生押韵违反的迅速探测和对韵律信息的初步整合。对于音乐训练组被试，相比于仅一种违反(如韵母违反，声调合适)，全违反条件(即韵母违反，声调违反) 下的整合难度较小，诱发的正波波幅相应更小；同时，相较于全合适条件(韵母合适，声调合适)，被试整合仅一种违反时(即韵母合适，声调违反或韵母违反，声调合适)消耗的认知资源更多，诱发的正波波幅也更大。相反，对照组在此时间窗口没有诱发任何显著的效应，可能表明普通人无法在绝句末迅速完成对绝句内出现的任何押韵违反的探测和整合。

不过，对照组在绝句末也会对绝句内发生的押韵违反进行深入分析，体现为声调和韵母合适时，韵母和声调不一致会诱发一个 300 ms 开始出现的类似于 N400 的负效应。不管该效应反映了语义(Magne et al., 2007; Marie et al., 2011; Perrin & García-Larrea, 2003)还是韵律(Chen et al., 2016; Li et al., 2014)的整合加工，当前结果与以往考察声调和韵母即时加工的研究诱发了相似的效应(Schirmer et al., 2005; Hu et al., 2012; Li et al., 2014)。在此时间窗口，音乐训练组仅在韵母和声调违反时才会继续对声调、韵母一致性进行深入分析。这表明音乐训

练组在上一阶段已经基本完成了韵律信息的整合；在此时间窗对声调和韵母违反的深入分析则体现了他们对韵律信息的感知加工更为精细。研究发现长期的音乐训练可以促进工作记忆广度和刷新能力的提升(Nutley, Darki, & Klingberg, 2014; Slevc Davey, Buschkuehl, & Jaeggi, 2016)，包括言语工作记忆(Clayton et al., 2016; Hansen, Wallentin, & Vuust, 2013; Roden, Kreutz, & Bongard, 2012)。为了完成当前押韵判断任务，听者需要动态存储诗句押韵信息并且在诗句末提取相关信息做出行为反应。音乐训练组可能相比对照组能更有效调动储存在工作记忆的信息，从而完成对声调和韵母更为精细和深入的加工过程。

以往采用被动 oddball 范式，探究音乐训练影响前注意阶段声调加工的研究发现，音乐训练组在声调变化时相比普通人会诱发更大的 MMN、P3a 成分(Nan et al., 2018; Tang, Xiong, Zhang, Dong, & Nan, 2016)。当前研究中，我们仅在音乐训练组发现声调违反条件下的反应时显著短于声调合适条件；即使韵母违反，音乐训练组仍会深入加工声调的合适性，体现在早期的正效应和晚期的负效应上。这可能表明在有意识注意状态下音乐训练组对声调信息仍然更为敏感。以往研究发现，不论音乐训练者的母语为声调还是非声调语言，其声调识别能力均优于非音乐训练者(Delogu, Lampis, & Belardinelli, 2010; Gottfried, Staby, & Ziemer, 2004; Cooper & Wang, 2012; Tang et al., 2016)。本研究结果则进一步表明，无论是在即时的声调加工过程中，还是在延迟的整合加工时，音乐训练者均相比普通人对声调投入更多注意资源，也更为敏感。

OPERA 理论指出音乐训练使得音乐训练者在基础声学线索的感知、一般认知能力等方面有很大的提高，并且伴随脑结构和功能的变化，为音乐能力在语言领域的迁移提供基础。本研究发现长期的音乐训练影响了人们对诗句韵律整合加工，使得音乐训练组在脑皮层水平上表现出了更加迅速、精细的神经整合过程。音乐训练在语言领域的这一迁移效应支持了 OPERA 理论，并且将现有探究音乐训练促进言语韵律信息即时加工的问题拓展到了对整体韵律信息的整合加工中。此外，本研究在有意识的主动注意状态下发现母语为声调语言的音乐训练组仍具加工优势，表明音乐训练的影响广泛，不仅体现在简单的声音感知方面，还体现在更加复杂、高级的认知加工过程。

5 结论

本研究采用 ERP 技术，通过让被试完成七言绝句的押韵判断任务，考察音乐训练组和对照组在诗歌末完成韵律信息的整合过程。结果发现，虽然两类被试都在声调/韵母合适的情况下，引起了由韵母/声调违反导致的整合困难，但是整合阶段不同：音乐训练组在早期就开始并快速完成整合，对照组则在较晚期时间窗才开始。此外，音乐训练组从早期就表现出对声调信息的敏感性，表现为即使韵母违反，声调信息的一致性仍会影响整合过程且贯穿整个加工过程；在晚期整合阶段，声调违反时，韵母违反也会增加音乐训练组的整合难度。综上，相比于对照组，音乐训练组对韵律违反信息(尤其是声调)的加工更敏感和快速，且对不同的违反类型有更精细的差异性反应。

参 考 文 献

Arthur, W., & Day, D. V. (1994). Development of a short form for the raven advanced progressive matrices test. *Educational & Psychological Measurement, 54*(2), 394–403.

Atienza, M., Cantero, J. L., & Dominguez-Marin, E. (2002). The time course of neural changes underlying auditory perceptual learning. *Learning & Memory*, *9*(3), 138–150.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*(2), 425–434.

Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience, 40*(4), 2662–2673.

Bohan, J., Leuthold, H., Hijikata, Y., & Sanford, A. J. (2012). The processing of good-fit semantic anomalies: An ERP investigation. *Neuropsychologia, 50*(14), 3174–3184.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56*(1), 103–128.

Carpenter, M., Cranford, J. L., Hymel, M. R., de Chicchis, A. R., & Holbert, D. (2002). Electrophysiologic signs of attention versus distraction in a binaural listening task. *Journal of Clinical Neurophysiology*, *19*(1), 55–60.

Chen, J., Liu, L., Wang, R., & Shen, H. Z. (2017). The effect of musical training on executive functions. *Advances in Psychological Science, 25*(11), 1854–1864.

[陈杰, 刘雷, 王蓉, 沈海洲. (2017). 音乐训练对执行功能的影响. *心理科学进展, 25*(11), 1854–1864.]

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844–2854.

Chen, Q. R., Zhang, J. J., Xu, X. D., Scheepers, C., Yang, Y. M., & Tanenhaus, M. K. (2016). Prosodic expectations in silent reading: ERP evidence from rhyme scheme and semantic congruence in classic Chinese poems. *Cognition, 154*, 11–21.

Chobert, J., François, C., Velay, J.-L., & Besson, M. (2012). Twelve months of active musical training in 8-to 10-year-

old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967.

Clayton, K. K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A. D., & Kidd, G. (2016). Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PloS One*, *11*(7), e0157638.

Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience, 17*(1), 168–182.

Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on cantonese word learning. *The Journal of the Acoustical Society of America*, *131*(6), 4756–4769.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732–744.

Dahan, D. (2015). Prosody and language comprehension. *Wiley Interdisciplinary Reviews Cognitive Science, 6*(5), 441–452.

de Vincenzi, M., Job, R., di Matteo, R., Angrilli, A., Penolazzi, B., Ciccarelli, L., & Vespignani, F. (2003). Differences in the perception and time course of syntactic and semantic violations. *Brain and Language*, *85*(2), 280–296.

Delogu, F., Lampis, G., & Belardinelli, M. O. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology*, *22*(1), 46–61.

Dilley, L. C., Mattys, S. L., & Vinke, L. (2010). Potent prosody: Comparing the effects of distal prosody, proximal prosody, and semantic context on word segmentation. *Journal of Memory and Language, 63*(3), 274–294.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self–paced reading methodologies. *Psychophysiology, 44*(6), 927–935.

Fan, W., Zhong, Y. P., Li, J., Yang, Z. L., Zhan, Y. L., Cai, R. H., & Fu, X. L. (2016). Negative emotion weakens the degree of self-reference effect: Evidence from ERPs. *Frontiers in Psychology, 7*, 1408.

François, C., Chobert, J., Besson, M., & Schön, D. (2012). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043.

Franklin, M. S., Sledge Moore, K., Yip, C.-Y., Jonides, J., Rattray, K., & Moher, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, *36*(3), 353–365.

Gottfried, T. L., Staby, A. M., & Ziemer, C. J. (2004). Musical experience and mandarin tone discrimination and imitation. *The Journal of the Acoustical Society of America*, *115*(5), 2545–2545.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15*(6), 883–899.

Hansen, M., Wallentin, M., & Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychology of Music*, *41*(6), 779–793.

Hausen, M., Torppa, R., Salmela, V. R., Vainio, M., & Särkämö, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology, 4*(2), 566.

Hu, J., Gao, S., Ma, W., & Yao, D. (2012). Dissociation of tone and vowel processing in Mandarin idioms. *Psychophysiology, 49*(9), 1179–1190.

Huang, X. J., Yang, J.-C., Zhang, Q., & Guo, C. Y. (2014). The time course of spoken word recognition in mandarin Chinese: A unimodal ERP study. *Neuropsychologia*, *63*, 165–174.

Huang, X., Liu, X., Yang, J. C., Zhao, Q., & Zhou, J. (2018). Tonal and vowel information processing in Chinese spoken word recognition: An event-related potential study. *Neuroreport, 29*(5), 1.

James, C. E., Oechslin, M. S., van de Ville, D., Hauert, C.-A., Descloux, C., & Lazeyras, F. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Structure and Function*, *219*(1), 353–366.

Jang, C. M. (2016). *Psychology of music*. Shanghai, China: East China Normal University Press.

[蒋存梅. (2016). *音乐心理学*. 上海: 华东师范大学出版社.]

Jin, H., Zhong, W. F., Xu, G. P., Cai, M. X., Yang, Y. F., Mo, L. (2009). The time course of world knowledge integration in sentence comprehension. *Acta Psychologica Sinica, 41*(7), 565–571.

[金花, 钟伟芳, 徐贵平, 蔡梦娴, 杨玉芳, 莫雷. (2009). 世界知识在句子理解中的整合时程. *心理学报, 41*(7), 565–571.]

Julia, H., Caroline, W., Caterina, P., Hubert, T., Barbara, H., & Isabell, W. (2013). Brain response to prosodic boundary cues depends on boundary position. *Frontiers in Psychology, 4*, 421.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*(5), 770–814.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of experimental psychology: General*, *111*(2), 228–238.

Kotz, S. A., Meyer, M., Alter, K., Besson, M., von Cramon, D. Y., & Friederici, A. D. (2003). On the lateralization of emotional prosody: An event-related functional MR investigation. *Brain and Language, 86*(3), 366–376.

Kühnis, J., Elmer, S., Meyer, M., & Jäncke, L. (2013). The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: An EEG study. *Neuropsychologia, 51*(8), 1608–1618.

Kuperberg, G. R., Kreher, D. A., Goff, D., McGuire, P. K., & David, A. S. (2006). Building up linguistic context in schizophrenia: Evidence from self-paced reading. *Neuropsychology, 20*(4), 442–452.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of psychology*, *62*, 621–647.

Lee, C.-Y., & Hung, T.-H. (2008). Identification of mandarin tones by English-speaking musicians and nonmusicians. *The Journal of the Acoustical Society of America*, *124*(5), 3235–3248.

Li, W. J., Liu, M., Zhang, Z. H., Deng, N. L., & Xing, Y. S. (2018). Neural processing of ambiguous Chinese phrases of stutters. *Acta Psychologica Sinica, 50*(12), 1323–1335.

[李卫君, 刘梦, 张政华, 邓娜丽, 邢钰珊. (2018). 口吃者加工汉语歧义短语的神经过程. *心理学报, 50*(12), 1323–1335.]

Li, W. J., Wang, L., & Yang, Y. F. (2014). Chinese tone and vowel processing exhibits distinctive temporal characteristics: An electrophysiological perspective from classical Chinese poem processing. *PLOS ONE, 9*(1), e85683.

Li, W. J., Yang, Y. F. (2010). The cognitive processing of prosodic boundary and its related brain effect in quatrain. *Acta Psychologica Sinica, 42*(11), 1021–1032.

[李卫君, 杨玉芳. (2010). 绝句韵律边界的认知加工及其脑电效应. *心理学报, 42*(11), 1021–1032.]

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain*, *133*(6), 1682–1693.

Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic

lengthening on semantic processing in spoken French: Behavioral and electrophysiological evidence. *Cerebral Cortex*, *17*(11), 2659–2668.
Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211.
Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*(2), 294–305.
Marques, C., Moreno, S., Luís Castro, S., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463.
Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research*, *1194*, 81–89.
Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research, 1228*, 161–176.
Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research, 308*, 84–97.
Musacchia, G., Strait, D., & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, *241*(1–2), 34–42.
Nan, Y. (2017). The facilitation effect of music learning on speech processing. *Advances in Psychological Science, 25*(11), 1844–1853.
[南云. (2017). 音乐学习对语言加工的促进作用. *心理科学进展, 25*(11), 1844–1853.]
Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., ... Desimone, R. (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proceedings of the National Academy of Sciences*, *115*(28), E6630–E6639.
Nan, Y., Sun, Y. N., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain*, *133*(9), 2635–2642.
Nutley, S. B., Darki, F., & Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Frontiers in Human Neuroscience, 7*, 926.
Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes, 8*(4), 413–437.
Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*(6), 739–773.
Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes, 14*(3), 283–317.
Partanen, E., Vainio, M., Kujala, T., & Huotilainen, M. (2011). Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles. P*sychophysiology, 48*(10), 1372–1380.
Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681.
Patel, A. D., (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*. *308*, 98–108.
Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, *87*(1), B35–B45.
Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *The Journal of the Acoustical Society of America*, *119*(5), 3034–3047.
Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language*, *61*(1), 123–144.
Perrin, F., & García-Larrea, L. (2003). Modulation of the N400 potential during auditory phonological/semantic interaction. *Cognitive Brain Research*, 17(1), 36–47.
Praamstra, P., & Stegeman, D. F. (1993). Phonological effects on the auditory N400 event-related brain potential. *Cognitive Brain Research*, *1*(2), 73–86.
Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(4), 1061–1080.
Reinke, K. S., He, Y., Wang, C. H., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*(3), 781–791.
Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: a longitudinal study. *Frontiers in Neuroscience*, *3*, 572.
Schirmer, A., Tang, S. L., Penney, T. B., Gunter, T. C., & Chen, H. C. (2005). Brain responses to segmentally and tonally induced semantic violations in Cantonese. *Journal of Cognitive Neuroscience, 17*(1), 1–12.
Schön, D., Besson, M., & Magne, C. L. (2004). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341–349.
Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience, 23*(13), 5545–5552.
Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *NeuroReport*, *16*(16), 1781–1785.
Slevc, L. R., Davey, N. S., Buschkuehl, M., & Jaeggi, S. M. (2016). Tuning the mind: Exploring the connections between musical ability and executive functions. *Cognition, 152*, 199–211.
Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early musical training and white-matter plasticity in the corpus callosum: Evidence for a sensitive period. *Journal of Neuroscience, 33*(3), 1282–1290.
Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*(2), 191–196.
Tang, W., Xiong, W., Zhang, Y.-X., Dong, Q., & Nan, Y. (2016). Musical experience facilitates lexical tone processing among mandarin speakers: Behavioral and neural evidence. *Neuropsychologia, 91*, 247–253.
Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in mandarin Chinese. *Language and Cognitive Processes, 23*(5), 689–708.
Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear and hearing*, *22*(2), 79–90.
Wang, X., Ossher, L., & Reuter-Lorenz, P. A. (2015). Examining the relationship between skilled music training and attention. *Consciousness & Cognition, 36*, 169–179.
Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*(4), 420–422.

Wu, H., Ma, X. H., Zhang, L. J., Liu, Y. Y., Yang, Z., & Hua, S. (2015). Musical experience modulates categorical perception of lexical tones in native chinese speakers. *Frontiers in Psychology, 6*, 436.

Wu, M. J., & Zhu, H. D. (2001). *Chinese rhythmology*. Beijing, China: Language & Culture Press.
[吴洁敏, 朱宏达. (2001). *汉语节律学*. 北京:语文出版社.]

Zioga, I., Luft, C. D. B., & Bhattacharya, J. (2016). Musical training shapes neural responses to melodic and prosodic expectation. *Brain Research*, *1650*, 267–282.

Musical training improves rhythm integrative processing of classical Chinese poem

ZHANG Zhenghua[1,2]; HAN Mei[1,2]; ZHANG Fang[3]; LI Weijun[1,2]

([1] *Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, Dalian* 116029, *China)*

([2] *Key Laboratory of Brain and Cognitive Neuroscience, Liaoning Province, Dalian* 116029, *China*)

([3] *Conservatory of Music, Liaoning Normal University, Dalian* 116029, *China*)

Abstract

Long-term rigorous musical training changes the brain structure and function, which impacts speech processing. Many studies have demonstrated that subcortical and cortical neural plasticity interact to give musicians linguistic advantages, allowing them to construct more elaborate perceptions of speech signals than non-musicians. There is ample empirical support for integrative processing at the end of a sentence. Studies have found that participants completed integrative processing of syntactic or semantic information at the end of a sentence containing an earlier syntactic or semantic violation. As speech rhythms play an important role in establishing the lexical, syntactic, and semantic representation of sentences, is it necessary for listeners to integrate rhythm at the end of sentence containing an earlier rhythm violation? Further, it remains unknown whether and (if yes) how musical training contributes to this process. The present ERP study used the rhyming judgment task to explore these questions using Chinese poems.

We used 160 unfamiliar, seven-character quatrains as experimental materials, with the last syllable (that is, targets) in the second line of the quatrain either appropriate or deviating from the appropriate syllable in tone, vowel, or both. Fifty undergraduates participated in this experiment. A total of 25 were musicians (9 males, mean age 20.56) with more than 10 years of formal music training who started at 7 years old or younger. The other 25 were non-musicians (9 males, mean age 20.84) who had never received any formal music training. Both were told to listen carefully and complete a ‘‘yes’’ or ‘‘no’’ judgment task on the appropriateness of the metrical feet. The EEG was recorded using 64 electrodes placed according to the international extended 10-20 system. The continuous EEG data were segmented into epochs of -200 to 1000 ms relative to the onset of the last syllable of each quatrain. The ERPs were analyzed and compared between conditions at time windows of 100~300 ms and 300~750 ms after the onset of the last syllable.

The behavioral results indicated that average accuracy rate in rhyming judgment task for musicians (83.24%, *SE* = 6.64%) was not significantly different from nonmusicians (82.64%, *SE* = 9.15%), $t(48) = 0.27$, $p = 0.791$. However, musicians responded faster when the tone was inappropriate compared to the appropriate condition, $F(1, 48) = 16.88$, $p < 0.001$, $\eta_p^2 = 0.26$. ERP results indicated that during 100–300 ms the interaction between Subject, Vowel and Tone was significant in the midline, $F(1, 48) = 7.59$, $p = 0.008$, $\eta_p^2 = 0.14$, and lateral areas, $F(1, 48) = 7.54$, $p = 0.008$, $\eta_p^2 = 0.14$. Simple effect analysis indicated that for musicians, the inappropriate vowel elicited larger positivities than the appropriate vowel when the tone was appropriate [midline: $F(1, 48) = 9.84$, $p = 0.003$, $\eta_p^2 = 0.17$; lateral: $F(1, 48) = 15.41$, $p < 0.001$, $\eta_p^2 = 0.24$]; the inappropriate tone elicited larger positivities than the appropriate tone when the vowel was appropriate[midline: $F(1, 48) = 3.80$, $p = 0.057$, $\eta_p^2 = 0.07$; lateral: $F(1, 48) = 7.27$, $p = 0.010$, $\eta_p^2 = 0.13$]. Besides, the inappropriate tone elicited smaller positivities than the appropriate tone when the vowel was inappropriate[midline: $F(1, 48) = 10.68$, $p = 0.002$, $\eta_p^2 = 0.18$; lateral: $F(1, 48) = 7.37$, $p = 0.009$, $\eta_p^2 = 0.13$]. However, for nonmusicians, no significant difference was found in this time window (*ps* > 0.1). During 300–750 ms, we found that the interaction between Subject, Vowel and

Tone was significant in the midline, $F(1, 48) = 6.66$, $p = 0.013$, $\eta_p^2 = 0.12$, and lateral areas, $F(1, 48) = 5.60$, $p = 0.022$, $\eta_p^2 = 0.10$. Simple effect analysis indicated that for musicians, the inappropriate vowel elicited larger negativities than the appropriate vowel when the tone was inappropriate [midline: $F(1, 48) = 5.14$, $p = 0.028$, $\eta_p^2 = 0.10$; lateral: $F(1, 48) = 2.92$, $p = 0.094$, $\eta_p^2 = 0.06$], and the inappropriate tone elicited larger negativities than the appropriate tone when the vowel was inappropriate[midline: $F(1, 48) = 12.94$, $p = 0.001$, $\eta_p^2 = 0.21$; lateral: $F(1, 48) = 9.65$, $p = 0.003$, $\eta_p^2 = 0.17$]. However, for nonmusicians, the inappropriate vowel elicited larger negativities than the appropriate vowel when the tone was appropriate [midline: $F(1, 48) = 15.07$, $p < 0.001$, $\eta_p^2 = 0.24$; lateral: $F(1, 48) = 12.04$, $p = 0.001$, $\eta_p^2 = 0.20$], while the inappropriate tone elicited larger negativities than the appropriate tone when the vowel was appropriate[midline: $F(1, 48) = 10.88$, $p = 0.002$, $\eta_p^2 = 0.19$; lateral: $F(1, 48) = 8.27$, $p = 0.006$, $\eta_p^2 = 0.15$].

In summary, the current study found that both musicians and non-musicians completed rhythm integrative processing at the end of sentence containing an earlier rhythm violation that was reflected by a late negative effect (during 300~750 ms). More importantly, through years of musical training, the musicians were more sensitive and faster at the integrative processing of rhythm information as reflected by an early positive component during 100~300 ms. In line with the behavioral data, the musicians developed increased sensitivity to tone variations during both early and late time windows. Taken together, both musicians and non-musicians completed the integrative processing of rhythm information at the end of quatrains, whereas the musicians were more sensitive and faster.

Key words musical training; rhythm; poem; integrative processing